



# Full wwPDB NMR Structure Validation Report ⓘ

Mar 6, 2022 – 05:00 PM EST

PDB ID : 2PX9  
Title : The intrinsic affinity between E2 and the Cys domain of E1 in Ubiquitin-like modifications  
Authors : Wang, J.H.; Hu, W.D.; Cai, S.; Lee, B.; Song, J.; Chen, Y.  
Deposited on : 2007-05-14

This is a Full wwPDB NMR Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/NMRValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
RCI : v\_1n\_11\_5\_13\_A (Berjanski et al., 2005)  
PANAV : Wang et al. (2010)  
ShiftChecker : 2.27  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.27

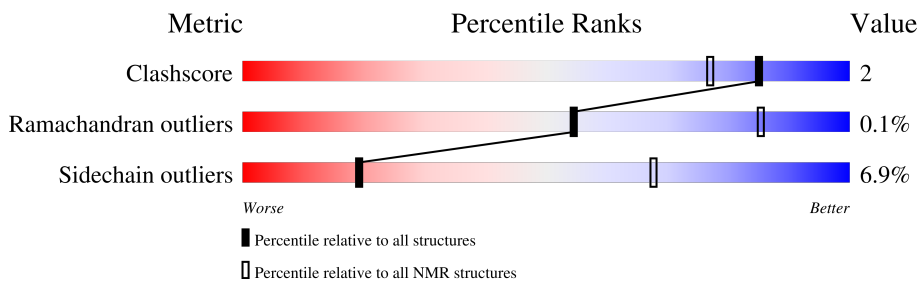
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment was not calculated.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and a grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	217	
2	B	158	

## 2 Ensemble composition and analysis i

This entry contains 20 models. Model 10 is the overall representative, medoid model (most similar to other models).

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	Residue range (total)	Backbone RMSD (Å)	Medoid model
1	A:166-A:213, A:239-A:291, A:304-A:382, B:1-B:158 (338)	0.50	10

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 3 clusters and 1 single-model cluster was found.

Cluster number	Models
1	1, 2, 3, 4, 5, 9, 10, 11, 12, 13, 14, 16, 20
2	6, 17, 18, 19
3	8, 15
Single-model clusters	7

### 3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5984 atoms, of which 2971 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called SUMO-activating enzyme subunit 2.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	217	3444	1093	1699	303	341	8	0

- Molecule 2 is a protein called SUMO-conjugating enzyme UBC9.

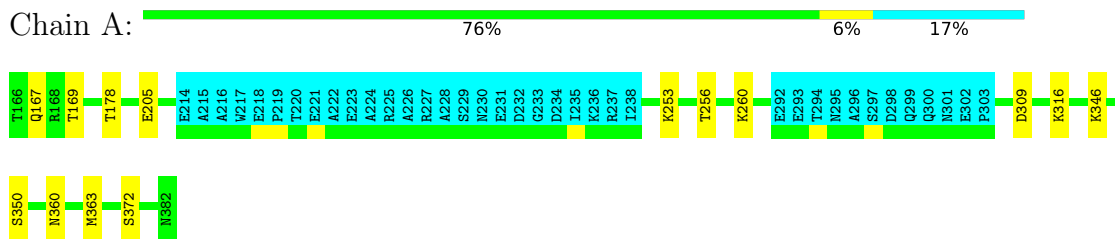
Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	158	2540	815	1272	218	227	8	0

## 4 Residue-property plots [i](#)

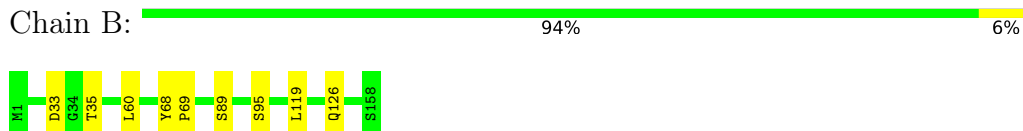
### 4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

- Molecule 1: SUMO-activating enzyme subunit 2



- Molecule 2: SUMO-conjugating enzyme UBC9

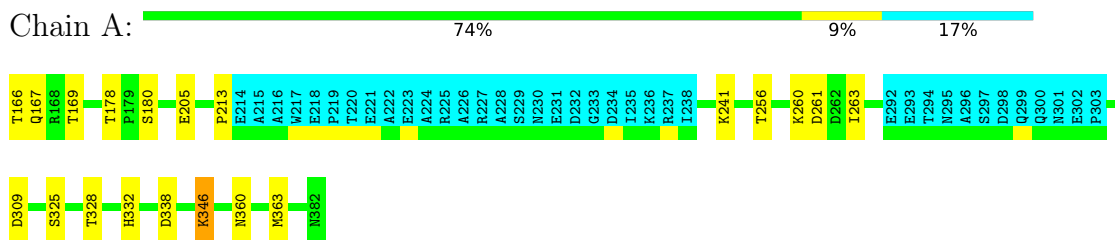


### 4.2 Scores per residue for each member of the ensemble


Colouring as in section 4.1 above.

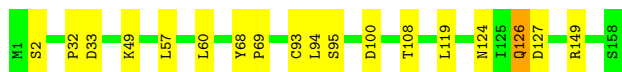
#### 4.2.1 Score per residue for model 1

- Molecule 1: SUMO-activating enzyme subunit 2



- Molecule 2: SUMO-conjugating enzyme UBC9

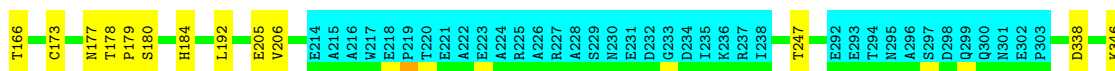
Chain B:  89% 11%




#### 4.2.2 Score per residue for model 2

- Molecule 1: SUMO-activating enzyme subunit 2

Chain A:  75% 8% 17%



- Molecule 2: SUMO-conjugating enzyme UBC9

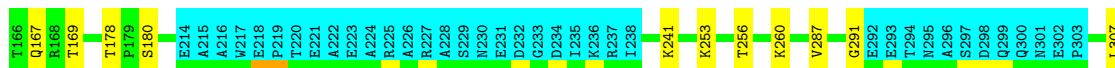
Chain B:  89% 11%




#### 4.2.3 Score per residue for model 3

- Molecule 1: SUMO-activating enzyme subunit 2

Chain A:  71% 12% 17%



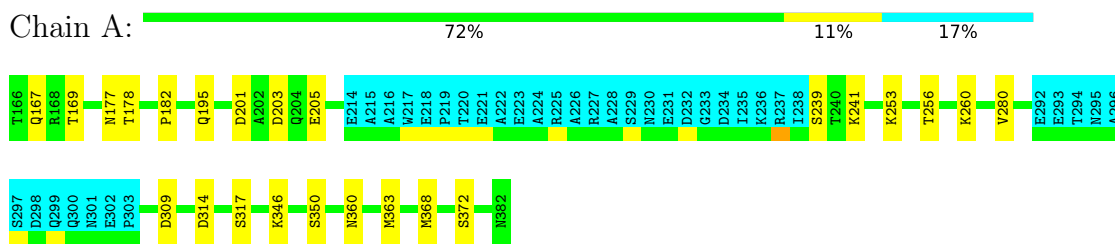
- Molecule 2: SUMO-conjugating enzyme UBC9

Chain B:  90% 9%

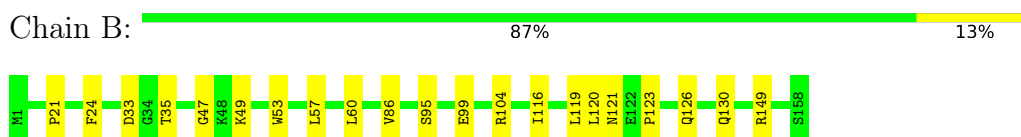


#### 4.2.4 Score per residue for model 4

- Molecule 1: SUMO-activating enzyme subunit 2

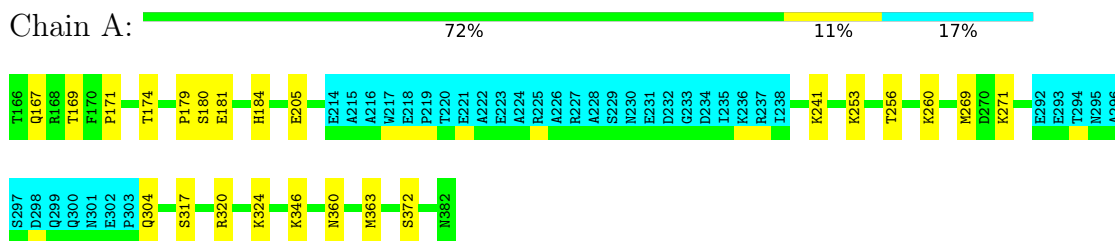


- Molecule 2: SUMO-conjugating enzyme UBC9

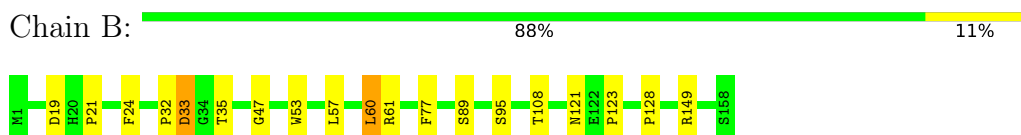


#### 4.2.5 Score per residue for model 5

- Molecule 1: SUMO-activating enzyme subunit 2

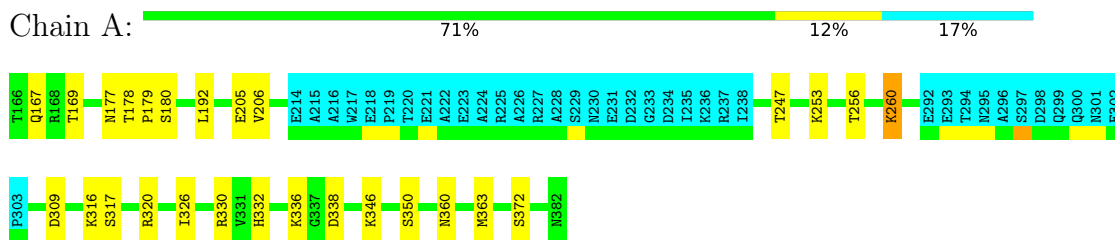


- Molecule 2: SUMO-conjugating enzyme UBC9




#### 4.2.6 Score per residue for model 6

- Molecule 1: SUMO-activating enzyme subunit 2



- Molecule 2: SUMO-conjugating enzyme UBC9

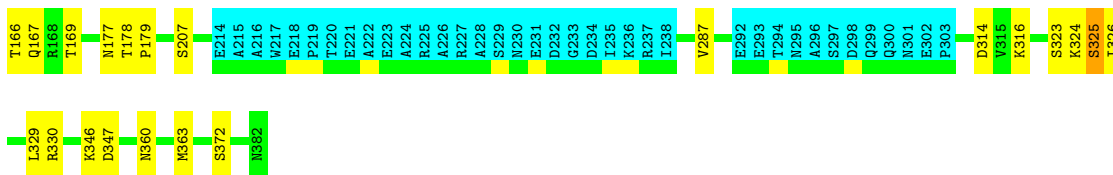
Chain B:  89% 11%




#### 4.2.7 Score per residue for model 7

- Molecule 1: SUMO-activating enzyme subunit 2

Chain A:  73% 9% 17%



- Molecule 2: SUMO-conjugating enzyme UBC9

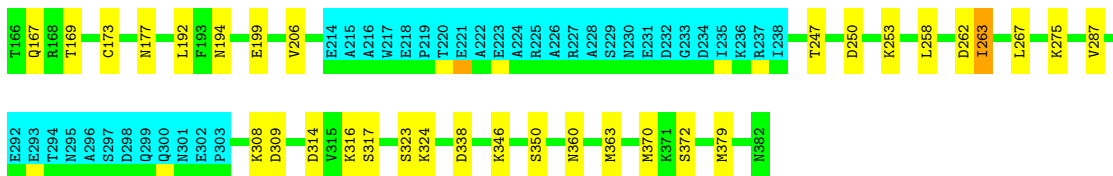
Chain B:  87% 13%




#### 4.2.8 Score per residue for model 8

- Molecule 1: SUMO-activating enzyme subunit 2

Chain A:  68% 14% 17%



- Molecule 2: SUMO-conjugating enzyme UBC9

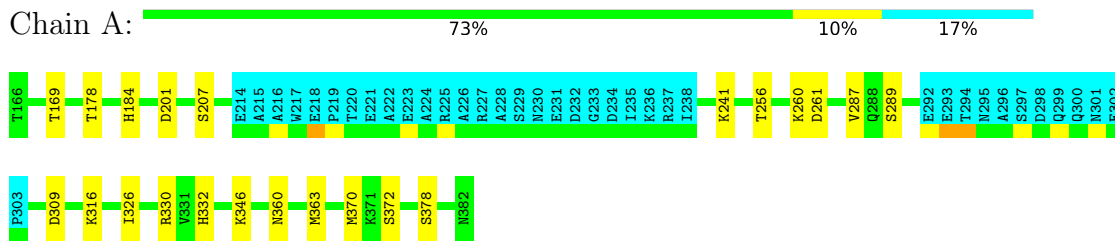
Chain B:  85% 15%



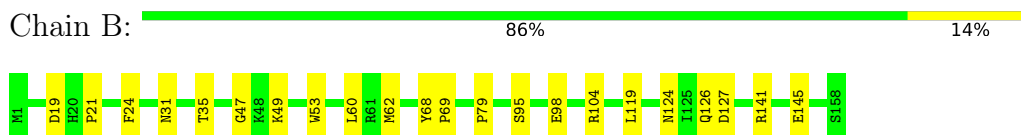
#### 4.2.9 Score per residue for model 9

- Molecule 1: SUMO-activating enzyme subunit 2



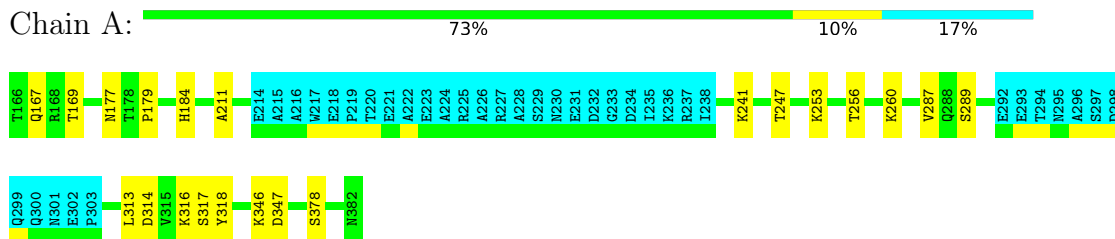


- Molecule 2: SUMO-conjugating enzyme UBC9

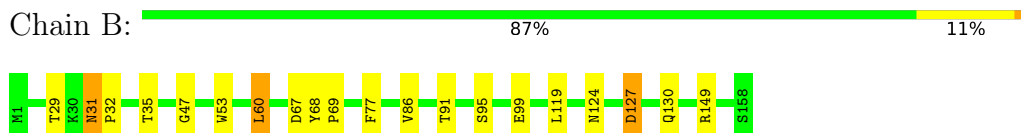


#### 4.2.10 Score per residue for model 10 (medoid)

- Molecule 1: SUMO-activating enzyme subunit 2

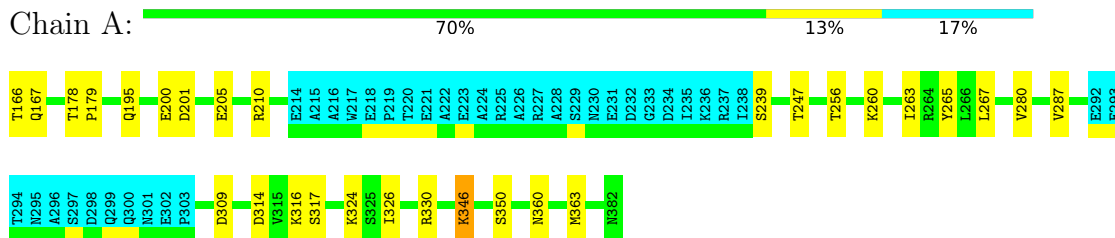


- Molecule 2: SUMO-conjugating enzyme UBC9

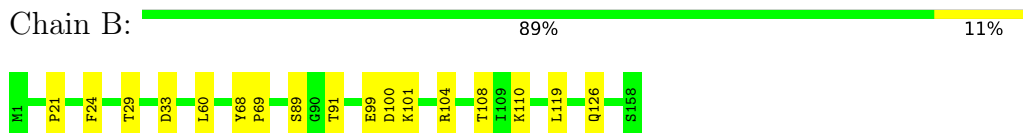


#### 4.2.11 Score per residue for model 11

- Molecule 1: SUMO-activating enzyme subunit 2

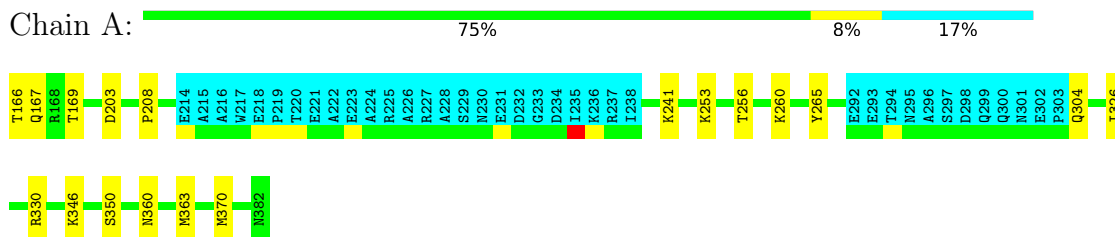


- Molecule 2: SUMO-conjugating enzyme UBC9

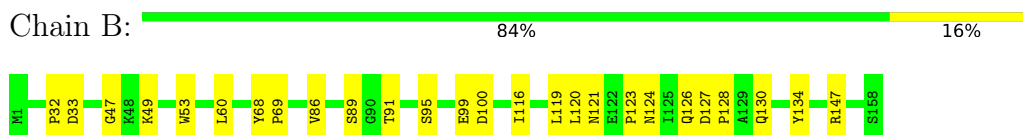


#### 4.2.12 Score per residue for model 12

- Molecule 1: SUMO-activating enzyme subunit 2

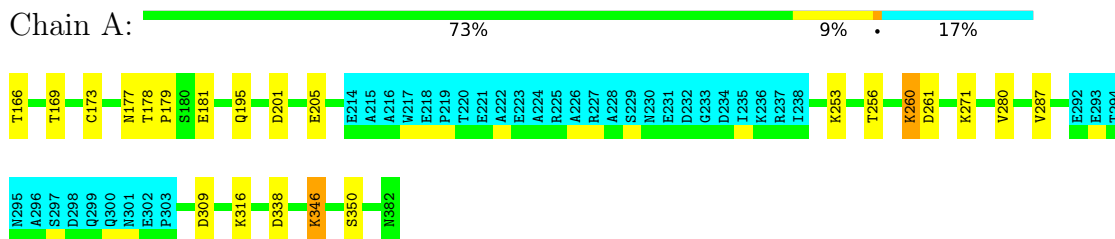


- Molecule 2: SUMO-conjugating enzyme UBC9

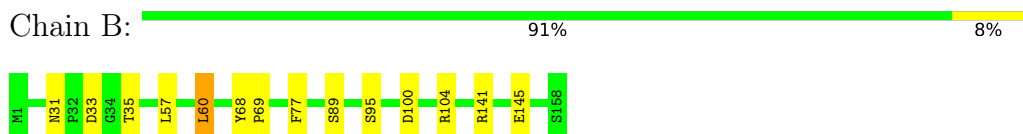


#### 4.2.13 Score per residue for model 13

- Molecule 1: SUMO-activating enzyme subunit 2

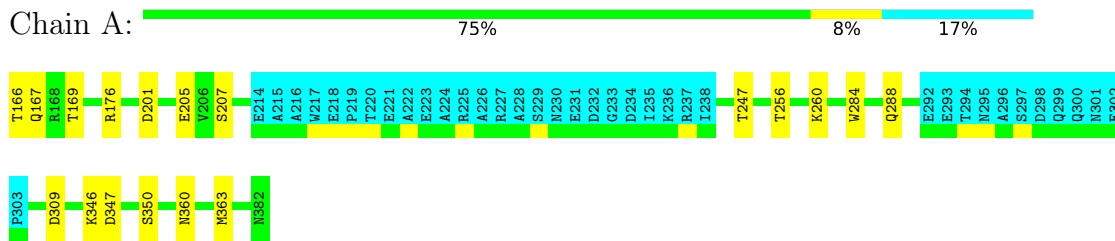


- Molecule 2: SUMO-conjugating enzyme UBC9

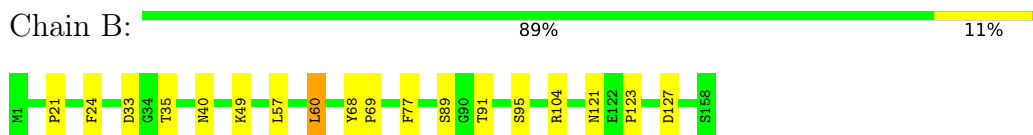


#### 4.2.14 Score per residue for model 14

- Molecule 1: SUMO-activating enzyme subunit 2

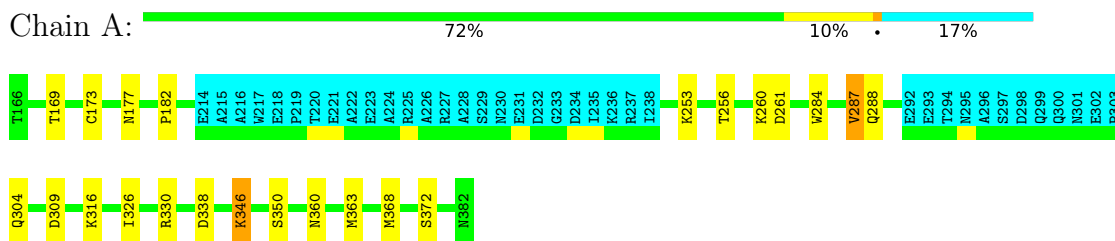


- Molecule 2: SUMO-conjugating enzyme UBC9

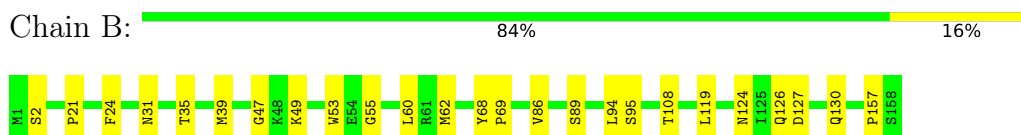


#### 4.2.15 Score per residue for model 15

- Molecule 1: SUMO-activating enzyme subunit 2

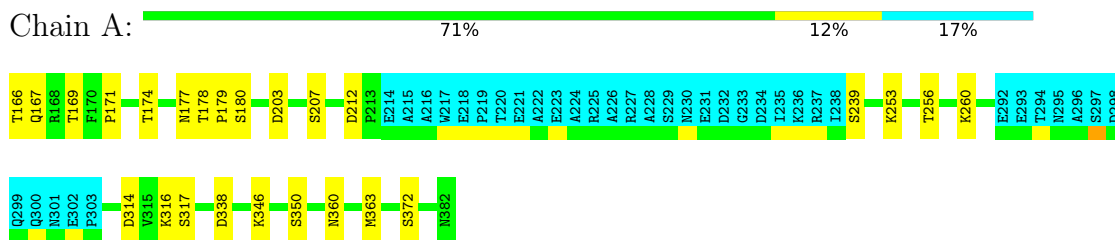


- Molecule 2: SUMO-conjugating enzyme UBC9




#### 4.2.16 Score per residue for model 16

- Molecule 1: SUMO-activating enzyme subunit 2



- Molecule 2: SUMO-conjugating enzyme UBC9

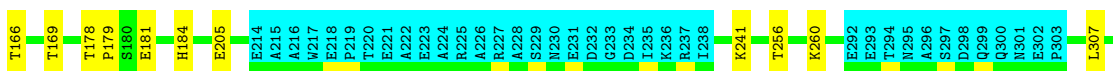
Chain B:  92% 8%



#### 4.2.17 Score per residue for model 17

- Molecule 1: SUMO-activating enzyme subunit 2

Chain A:  74% 9% 17%



- Molecule 2: SUMO-conjugating enzyme UBC9

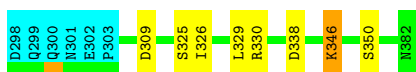
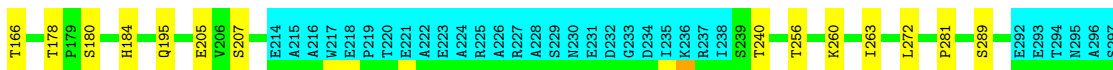
Chain B:  88% 11%



#### 4.2.18 Score per residue for model 18

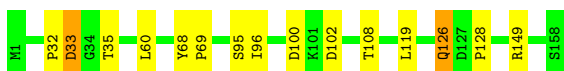
- Molecule 1: SUMO-activating enzyme subunit 2

Chain A:  73% 10% 17%



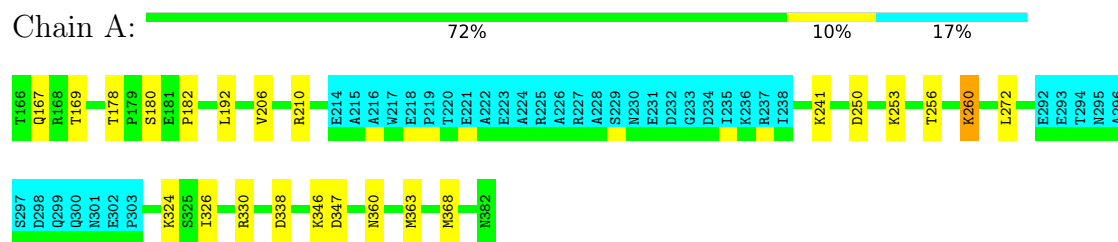
- Molecule 2: SUMO-conjugating enzyme UBC9

Chain B:  91% 8%

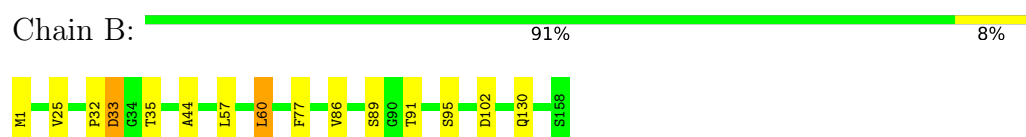


#### 4.2.19 Score per residue for model 19

- Molecule 1: SUMO-activating enzyme subunit 2

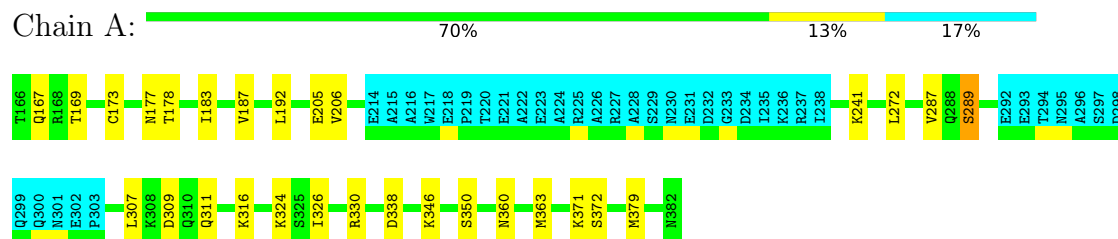


- Molecule 2: SUMO-conjugating enzyme UBC9

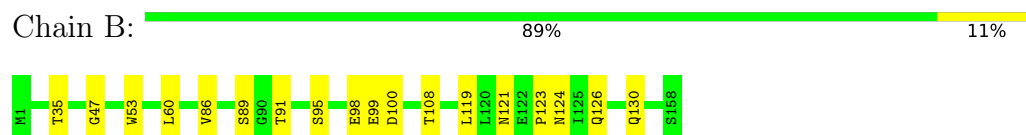


#### 4.2.20 Score per residue for model 20

- Molecule 1: SUMO-activating enzyme subunit 2



- Molecule 2: SUMO-conjugating enzyme UBC9



## 5 Refinement protocol and experimental data overview

Of the 1000 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
HADDOCK	refinement	

No chemical shift data was provided.

## 6 Model quality i

### 6.1 Standard geometry i

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 6.2 Too-close contacts i

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	1456	1445	1439	6±2
2	B	1268	1272	1270	6±1
All	All	54480	54340	54180	235

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:360:ASN:HA	1:A:363:MET:SD	0.66	2.29	4	17
2:B:96:ILE:HA	2:B:102:ASP:O	0.58	1.97	18	3
1:A:326:ILE:O	1:A:330:ARG:HG3	0.56	2.00	12	10
1:A:182:PRO:HA	1:A:368:MET:SD	0.55	2.41	19	3
1:A:195:GLN:NE2	1:A:205:GLU:HA	0.55	2.16	18	4
1:A:173:CYS:O	1:A:177:ASN:HB2	0.55	2.02	15	5
1:A:256:THR:O	1:A:260:LYS:HB3	0.55	2.02	5	16
2:B:68:TYR:CD1	2:B:69:PRO:HA	0.54	2.36	10	14
1:A:177:ASN:O	1:A:179:PRO:HD3	0.52	2.03	16	6
1:A:192:LEU:HD13	1:A:206:VAL:HG21	0.52	1.80	2	5
2:B:21:PRO:HB2	2:B:24:PHE:CD1	0.52	2.39	2	9
2:B:32:PRO:O	2:B:33:ASP:HB3	0.52	2.04	19	5
1:A:201:ASP:HA	2:B:104:ARG:NH1	0.52	2.20	9	4

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:250:ASP:OD2	1:A:253:LYS:HG2	0.51	2.05	19	2
1:A:176:ARG:HD3	1:A:205:GLU:O	0.51	2.05	14	1
2:B:121:ASN:O	2:B:123:PRO:HD3	0.50	2.07	4	8
1:A:258:LEU:HA	1:A:262:ASP:OD2	0.49	2.08	8	1
2:B:124:ASN:ND2	2:B:127:ASP:HB2	0.48	2.23	10	1
2:B:86:VAL:O	2:B:130:GLN:HB2	0.48	2.08	10	10
1:A:287:VAL:O	1:A:316:LYS:HG2	0.48	2.09	7	9
2:B:65:LYS:HE2	2:B:67:ASP:OD1	0.48	2.09	17	1
2:B:128:PRO:HB3	2:B:134:TYR:CG	0.48	2.44	12	1
2:B:68:TYR:CG	2:B:69:PRO:HA	0.47	2.44	1	8
1:A:363:MET:SD	1:A:370:MET:HA	0.47	2.50	8	3
1:A:363:MET:O	1:A:368:MET:HG2	0.47	2.09	3	1
2:B:124:ASN:HD21	2:B:126:GLN:NE2	0.47	2.07	1	1
2:B:124:ASN:ND2	2:B:126:GLN:HB3	0.47	2.24	15	5
2:B:31:ASN:HB3	2:B:35:THR:OG1	0.47	2.10	15	2
2:B:60:LEU:HB3	2:B:77:PHE:CD1	0.47	2.44	5	9
1:A:263:ILE:O	1:A:267:LEU:HG	0.46	2.10	8	2
1:A:346:LYS:HD3	1:A:346:LYS:H	0.46	1.69	13	4
2:B:47:GLY:HA3	2:B:53:TRP:O	0.46	2.10	4	10
1:A:317:SER:HA	1:A:320:ARG:NE	0.46	2.25	5	2
1:A:332:HIS:O	1:A:336:LYS:HG2	0.46	2.11	3	2
1:A:208:PRO:HB2	1:A:265:TYR:CD2	0.46	2.46	12	1
1:A:313:LEU:HB2	1:A:318:TYR:CE2	0.46	2.46	10	1
1:A:363:MET:HB2	1:A:368:MET:SD	0.45	2.50	19	2
2:B:126:GLN:O	2:B:128:PRO:HD3	0.45	2.11	18	4
1:A:284:TRP:O	1:A:288:GLN:HG2	0.45	2.12	14	2
2:B:32:PRO:O	2:B:33:ASP:HB2	0.45	2.12	3	3
1:A:263:ILE:HD11	1:A:281:PRO:HA	0.45	1.88	18	1
2:B:25:VAL:HB	2:B:44:ALA:HB3	0.45	1.89	19	1
1:A:171:PRO:O	1:A:174:THR:HG22	0.44	2.12	5	2
2:B:141:ARG:O	2:B:145:GLU:HG2	0.44	2.11	9	3
1:A:307:LEU:O	1:A:311:GLN:HG2	0.44	2.12	3	3
2:B:108:THR:HG22	2:B:110:LYS:H	0.44	1.73	11	1
1:A:371:LYS:HE2	1:A:379:MET:SD	0.44	2.53	20	1
1:A:210:ARG:HA	1:A:265:TYR:OH	0.43	2.13	11	1
2:B:100:ASP:OD2	2:B:101:LYS:HG3	0.43	2.12	11	1
1:A:284:TRP:O	1:A:287:VAL:HG12	0.43	2.13	15	1
1:A:325:SER:O	1:A:329:LEU:HG	0.43	2.12	18	2
1:A:194:ASN:HB3	1:A:199:GLU:O	0.43	2.14	8	1
1:A:201:ASP:HA	2:B:104:ARG:CZ	0.42	2.44	14	1
2:B:85:ASN:ND2	2:B:127:ASP:HB2	0.42	2.29	8	1

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:353:ASP:OD1	1:A:373:ARG:HD2	0.42	2.14	3	1
2:B:149:ARG:HA	2:B:149:ARG:NE	0.41	2.30	5	2
2:B:94:LEU:HD13	2:B:119:LEU:HD11	0.41	1.92	1	2
1:A:360:ASN:OD1	1:A:370:MET:HB3	0.41	2.16	12	1
2:B:142:VAL:O	2:B:146:LYS:HG3	0.41	2.15	16	1
2:B:124:ASN:OD1	2:B:126:GLN:HB3	0.41	2.15	17	1
1:A:360:ASN:O	1:A:363:MET:HG2	0.41	2.16	14	1
2:B:60:LEU:HD13	2:B:60:LEU:N	0.41	2.31	16	1
2:B:116:ILE:O	2:B:120:LEU:HG	0.41	2.16	4	2
1:A:181:GLU:O	1:A:184:HIS:HB2	0.41	2.16	5	2
2:B:31:ASN:OD1	2:B:32:PRO:HD2	0.41	2.16	10	1
1:A:183:ILE:O	1:A:187:VAL:HG23	0.40	2.16	20	1
1:A:346:LYS:H	1:A:346:LYS:HD3	0.40	1.76	1	1
2:B:55:GLY:O	2:B:157:PRO:HD3	0.40	2.16	15	1

## 6.3 Torsion angles [i](#)

### 6.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the backbone conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	178/217 (82%)	169±1 (95±1%)	9±1 (5±1%)	0±0 (0±0%)	50	82
2	B	156/158 (99%)	150±2 (96±1%)	6±2 (4±1%)	0±0 (0±0%)	54	85
All	All	6680/7500 (89%)	6377 (95%)	293 (4%)	10 (0%)	54	85

All 7 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	179	PRO	3
1	A	180	SER	2
1	A	177	ASN	1
2	B	128	PRO	1
2	B	101	LYS	1
1	A	212	ASP	1
2	B	32	PRO	1

### 6.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the sidechain conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	160/189 (85%)	148±2 (93±1%)	12±2 (7±1%)	17	65
2	B	137/137 (100%)	128±1 (94±1%)	9±1 (6±1%)	21	70
All	All	5940/6520 (91%)	5528 (93%)	412 (7%)	19	68

All 77 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	346	LYS	20
2	B	60	LEU	20
1	A	169	THR	17
2	B	95	SER	15
1	A	167	GLN	14
1	A	178	THR	14
2	B	89	SER	13
1	A	350	SER	12
2	B	35	THR	12
1	A	309	ASP	11
1	A	372	SER	11
2	B	33	ASP	11
2	B	119	LEU	11
1	A	166	THR	10
1	A	241	LYS	10
1	A	338	ASP	10
1	A	253	LYS	9
2	B	91	THR	9
1	A	314	ASP	8
2	B	49	LYS	7
2	B	57	LEU	7
2	B	127	ASP	7
2	B	99	GLU	7
1	A	180	SER	6
1	A	205	GLU	6
2	B	108	THR	6
1	A	247	THR	6
1	A	317	SER	6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	324	LYS	6
2	B	31	ASN	5
1	A	347	ASP	5
1	A	207	SER	5
1	A	261	ASP	4
2	B	100	ASP	4
2	B	126	GLN	4
2	B	149	ARG	4
1	A	184	HIS	4
1	A	289	SER	4
2	B	2	SER	3
2	B	62	MET	3
1	A	203	ASP	3
1	A	239	SER	3
1	A	280	VAL	3
1	A	304	GLN	3
1	A	260	LYS	3
2	B	29	THR	3
2	B	67	ASP	3
2	B	102	ASP	3
1	A	323	SER	3
1	A	272	LEU	3
1	A	263	ILE	2
1	A	325	SER	2
1	A	332	HIS	2
2	B	39	MET	2
2	B	70	SER	2
1	A	271	LYS	2
2	B	19	ASP	2
1	A	378	SER	2
2	B	147	ARG	2
1	A	328	THR	1
2	B	66	ASP	1
1	A	269	MET	1
2	B	61	ARG	1
2	B	97	LEU	1
1	A	275	LYS	1
1	A	308	LYS	1
1	A	379	MET	1
2	B	79	PRO	1
1	A	200	GLU	1
1	A	181	GLU	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
2	B	40	ASN	1
1	A	287	VAL	1
1	A	316	LYS	1
2	B	71	SER	1
1	A	240	THR	1
2	B	1	MET	1
2	B	98	GLU	1

### 6.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 6.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 6.6 Ligand geometry [i](#)

There are no ligands in this entry.

### 6.7 Other polymers [i](#)

There are no such molecules in this entry.

### 6.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 7 Chemical shift validation

No chemical shift data were provided